Cloud Computing Using RAID Model

Jigisha Pandya*, Mousumi Sarkar#
Department of MCA
* #*G. H. Raisoni institute of management and research, Wagholi, Pune, MH, India*
[1]jigishadpandya@gmail.com
[2]mousumi.sarkar@raisoni.net

Abstract — **Cloud Computing is the recently emerging technology these days, Cloud computing is used for SAAS, PAAS and IASS.**
Cloud computing allows to store data and maintain at remote place. Primary advantage of using cloud computing is efficient and cost effective data storage. Example of Cloud Service Providers (CSP) is Amazon and Google. Data hosted on cloud server of Google or Amazon is available to wide variety of clients throughout the world. As the amount of data stored on cloud server increases so does the security concerns.
In this paper we discussed security issues in the Cloud and possible solutions of some the problems. As a part of this paper we introduced you solution for data availability problem using RAID architecture.

Keywords – **Cloud Computing, Security Issues, Cloud Architecture, Cloud and RAID.**

I. INTRODUCTION

One significant advantage of cloud computing is data accessibility over the internet. Computing and IT resources are encapsulated as services, hiding all the details of implementation, deployment, maintenance and administration [1]. Distributed computing using cloud enables data accessibility to user via internet. One of the core concerns of sharing the data over the cloud is no granular level of control.

In Absence of granular control in single cloud environment security issue occur with data. Data may not be available in single cloud environment. Unlimited access to cloud operator of data in term of read write and share. In this paper we will discuss and seek solutions for security issues in cloud computing.

Rest of the paper is organized as follows: We present the data security problem of cloud computing in the next section and then discuss the details of requirement. In Section III explains the basic idea about RAID technology. In Section IV We have explain architecture of RAID technology Section V is for conclusion and Section VI is for References.

A. SECURITY PROBLEM OF CLOUD COMPUTING AND REQUIREMENT

Distributed computing and lack of granular level control on data may lead to security concern in cloud computing.

Following points elaborates on security concerns in data availability and security.

II. DATA AVAILABILITY AND APPLICATION FAILURE

In case of unavailability of services by single cloud provider data/application down time is unavoidable.

B. BLACK BOX APPROACH TO DATA POSITIONING

Location visibility of data is not known to end user i.e. geographical or physical. For example sometimes user might not know about country where data is stored. Cautious decision needs to be taken by data provider confirming guideline/ policies pertaining to data sensitive information.

C. CONTINGENCY PLAN FOR DATA CORRUPTION

In case of data corruption due to business transaction or physical device failure data provider can efficiently resolved the situation if database is available locally, where as data is stored in cloud resolution has to be done via established channels making process significantly slower and less efficient.

D. SECURITY OF DATA

Data provider need to check about security of their data for example data should be ciphered at all the level and ciphered schema should be robust.

E. DATA CONFIDENTIALITY

Confidentiality of data is also very important matter in cloud computing there should be some committee to decide high level privacy issues.

Customer should ensure that cloud provider should maintain all privacy and policies as desired.

III. PROMINENT RAID MODEL

Redundant array of independent disk is used to group several physical drives in array that can be defined in one or more logical drive. These logical drive(s) work as independent drive as operating system.

RAID 1: RAID 1 model creates mirror of data in two or more disks. This level of RAID model is decisive when data performance/ throughput is more important than data storage. RAID 1 model improves the data reliability by organizing data in two clusters (disks). Data is available and an access is faster due to separate copy of data in each disk and data can be address independently.
RAID 1 Model describe in Figure 1.

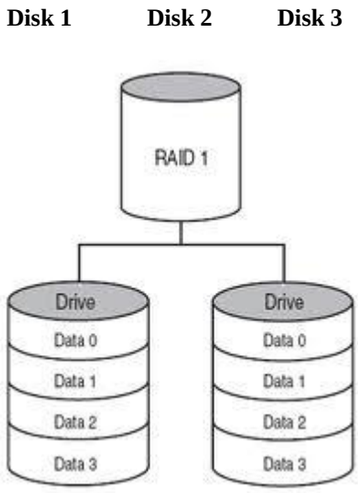

FIGURE 1: RAID1 MODEL

RAID 5: Distributed Parity level

RAID 5 uses a block level uses a block level striping with distributed parity. For an array of n drives, with $S_{\min}$ In RAID level 5 uses $S_{\min} \times (n-1)$.

Parity bit:

Series of data blocks are known as stripe. Parity bit crawls from one disk to another hence it is known as **Distributed Parity level.**

Parity bits are not read as they increase over head for disk controller.

Parity handling process:

- Read the old data block
- Read the old parity block
- Compare the old data block with the write request. For each bit that has flipped (changed from 0 to 1, or from 1 to 0) in the data block, flip the corresponding bit in the parity block
- Write the new data block
- Write the new parity block [2]

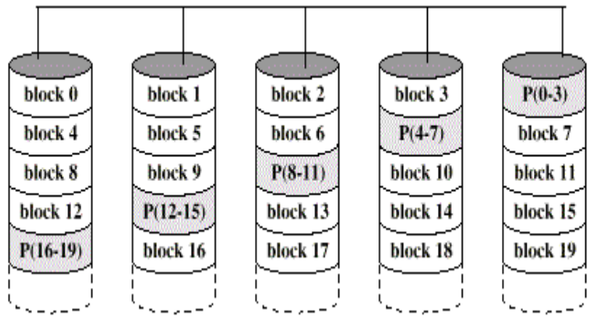

Figure 2

- Here P(16-19) contains parity for block 16, 17,18 and 19
- Similarly P(12-15) contains parity for block 12,13,14,15

IV. PROPOSED MODEL

We square measure proposing a model for cloud, supported RAID design. In our planned model we have used two model RAID 1 and RAID 5 level. In this configuration data availability is more; this configuration can sustain failure of all the disk in any array plus one disk of other array. This configuration is very accurate and used for fault tolerance, data availability and can also be used for data recovery in cloud system.

This combine architecture is known as RAID 51. In this architecture we combine RAID model 5 and RAID model1. RAID model one is used for making copy of the data or for that makes data available in disk failure .RAID 5 is used for striping of data using distributed parity bit this is very useful for data correction.

RAID 51: In this model we have created two RAID 5 mirror images that create mirror image for each other and even they don't have any idea that mirror image is exist. Mirroring is very useful for data availability as well as if any changes made in any disk that can easily traceable because of mirroring.

In our architecture RAID 5 is under RAID1 architecture using this architecture we get data correction because of RAID 5 parity data distribution across all the disks. As shown in figure 3.

This model is very accurate for data availability but this model is slow when multiple read write operation is performed. This is because of parity; parity must be updated at every write operation and it require to update in read modify and write sequence as well.

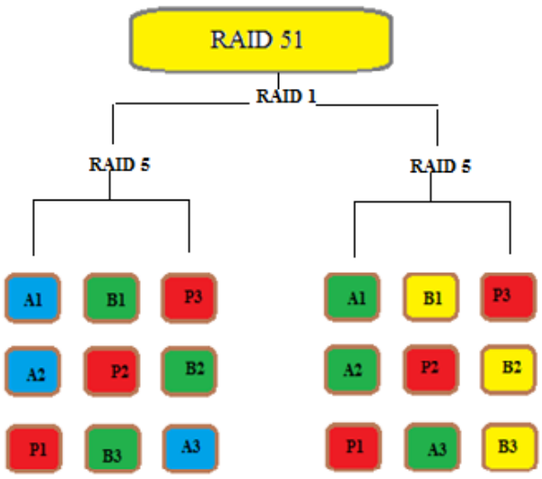

Figure 3

This performance issue can be little upgrade by using RAID 6 in place of RAID 5. It uses two parity block in each stripe. RAID 6 has advantage over RAID 5 because it does not require penalty for read operation but only performance penalty for write operation. This can increase performance greatly depend upon how RAID 6 implement storage architecture.

Calculate Parity:

Two parameters needs to calculate for allow data loss of any two drive. One of them is parity (P) which is XOR of data across stripe. A second parameter is Galois field GF (m) is introduced with m = 2k, where GF (m) $\cong F_2[x]/ (p(x))$ for irreducible polynomial p(x) of degree k.

D1,…Dn $\in$ GF(m) is data across different hard drives.

P and Q can be calculated using following formula

$$\mathbf{P = \bigoplus_i D_i = D_1 \oplus D2 \oplus \ldots \oplus Dn}$$

$$\mathbf{Q = \bigoplus_i g^i D_i = g^0 D_1 \oplus g^1 D2 \oplus \ldots \oplus g^n Dn}$$

Where $\oplus$ is a bitwise XOR operator and g^i is the action of a linear feedback shift register on a chunk of data. Thus, in the above formula, the calculation of P is just the XOR of each stripe. This is due to addition in any attribute two finite fields reduces to the XOR operation. The computation of Q is the XOR of a shifted operation of each stripe. Mathematically, the generator is an element of the field such that g^i is different for each nonnegative i satisfying i < n.

If one data drive is lost, the data can be recomputed from P just like with RAID 5. If two data drives are lost or the drive containing P is lost the data can be recovered from P and Q using a more complex process. Working out the details is not hard with field theory.

V. CONCLUSION

Cloud computing is cost and performance effective. There are some key issues in cloud computing for client as well as for vendors. Trusted computing, legal contract and cryptography is the main parameters for security of cloud computing. Even though for security issues no proper policy exist organizational data should be protect with proper policy. Client and vendor should clear about their security standards, services requirement, issues, termination rights and so on.

RAID one model is used as a protection against physical disk failure. It does not provide any protection against viruses or accidental change of file or deletion of data. This type of changes immediately mirrored in other disks. For example if data in one drive is damage due to some reason design of RAID 1 immediately copied (mirrored) this effect in other drives in array the same time. For this reason system using RAID 1 need to take a back up for updated data for restoring purpose. It would be self – evidence that system with redundancy also need reliable backup system.

The risk must be cautiously balanced against available safe guard. The suitable balance between the strength of controls and the relative risk associated with particular programs and operations must be ensured.

VI. REFERENCES

[1] Anil Gupta, Aaftab Qureshi, Parag Pande,"Cloud Computing Characteristics and Service Models: our own interpretation".
[2] Kresimir Popovic, et al., "Cloud" Computing issues and challanges" MIPRO 2010 May 24-28 Opatija, Croatia, pp 344-349.
[3] Gansen Zhao, et al., "Deployment Models: Towards Eliminating Security Concerns from Cloud Computing" IEEE 2010, pp 189-195.
[4] J. Dean and S. Ghemawat. "Mapreduce: simplified data processing on large clusters". *Commun. ACM*, 51(1):107–113, 2008.
[5] Apache Hadoop, 2009. http://hadoop.apache.org/.
[6] M. Isard, M. Budiu, Y. Yu, A. Birrell, and D. Fetterly."Dryad: distributed data-parallel programs from sequential building blocks". In EuroSys '07: Proceedings of the 2nd ACM SIGOPS/EuroSys European Conference on Computer Systems 2007, pages 59–72, New York, NY, USA, 2007. ACM.
[7] Condor DAGman, 2009. http://www.cs.wisc.edu/condor/dagman/. [18] D. Nurmi, R. Wolski, C. Grzegorczyk, G. Obertelli, S. Soman, L. Youseff, and D. Zagorodnov. "The eucalyptus open-source cloud-computing system". In Proceedings of Cloud Computing and Its Applications, October 2008.
[8] Nimbus. "Introduction to nimbus", 2009. http://workspace.globus.org/clouds/nimbus.html. [3] S. Beco, A. Maraschini, and F. Pacini. "Cloud computing and RESERVOIR project". NUOVO CIMENTO DELLA SOCIETA ITALIANA DI FISICA C-COLLOQUIA ON PHYSICS, 32(2), Mar-Apr 2009.
[9] CARMEN, 2009. http://www.carmen.org.uk/.

[10] Å. A. Nyre and M. G. Jaatun. "Privacy in a semantic cloud: What's trust got to do with it?". In The First International Conference on Cloud Computing, pages 107–118, 2009.
[11] S. Pearson, Y. Shen, and M. Mowbray. "A privacy manager for cloud computing". In The First international Conference on Cloud Computing, pages 90–106, 2009.
[12] L. Kaufman. "Data security in the world of cloud computing". IEEE SECURITY & PRIVACY, 7(4), July- August 2009.
[13] S. Creese, P. Hopkins, S. Pearson, and Y. Shen. "Data protection-aware design for cloud services". In The First International Conference on Cloud Computing, pages 119–130, 2009.
[14] L. Hu, S. Ying, X. Jia, and K. Zhao. "Towards an approach of semantic access control for cloud computing". In The First International Conference on Cloud Computing, pages 145–156, 2009.
[15] D. Chen, X. Huang, and X. Ren. "Access control of cloud service based on ucon". In The First International Conference on Cloud Computing, pages 559–564, 2009.
[16] T. Uemura, T. Dohi, and N. Kaio. "Availability analysis of a scalable intrusion tolerant architecture with two detection modes". In The First International Conference on Cloud Computing, pages 178–189, 2009.
[17] L. Yan, C. Rong, and G. Zhao. "Strengthen cloud computing security with federal identity management using hierarchical identity-based cryptography". In The First International Conference on Cloud Computing, pages 167–177, 2009.
[18] G. Zhao, J. Liu, Y. Tang, W. Sun, F. Zhang, X. ping Ye, and N. Tang. "Cloud computing: A statistics aspect of users". In The First International Conference on Cloud Computing, pages 347–358. Springer, 2009.
[19] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. "Above the clouds: A Berkeley view of cloud computing". Technical Report UCB/EECS- 2009-28, EECS Department, University of California, Berkeley, Feb 2009.
[20] M. Schumacher, E. Fernandez, D. Hybertson, and F. Buschmann. SECURITY PATTERNS: INTEGRATING SECURITY AND SYSTEMS ENGINEERING. John Wiley & Sons, 2005.
[21] T. Heyman, K. Yskout, R. Scandariato, and W. Joosen. "An analysis of the security patterns andscape". In SESS '07: Proceedings of the Third International Workshop on Software Engineering for Secure Systems, page 3, Washington, DC, USA, 2007. IEEE Computer Society.
[22] E. B. Fernandez, J. Wu, M. M. Larrondo-Petrie, and Y. Shao. On building secure scada systems using security patterns. In CSIIRW '09: Proceedings of the 5th Annual Workshop on Cyber Security and Information Intelligence Research, pages 1–4, New York, NY, USA, 2009. ACM.
[23] B. Blakley and C. Heath. SECURITY DESIGN PATTERNS, 2004. The Open Group Security Forum.
[24] K. J. Hughes. "Domain Based Security: enabling security at the level of applications and business processes", 2002. www.qinetiq.com.
[25] A. Singh, M. Srivatsa, and L. Liu. "Search-as-a-Service: Outsourced Search over Outsourced Storage". *ACM TRANSACTIONS ON THE WEB*, 3(4), September 2009.
[26] Wikipedia" www.wikipedia.org".

AUTHORS

Jigisha Pandya was born in 1980. She received M.C.A. degree from Saurashtra University. She joined as a lecturer in Bachelor of Computer Applications in H.J. Doshi college in 2003. Presently she is working as a Assistant Professor of Master of Computer Applications in G.H. Raisoni Institute of Mgmt. and Research. She has published 4 international papers in journals and conferences. Currently he is working on Cloud computing.

Mousumi Sarkar was born in 1980. She received M.C.A. degree from Birla Institute of Technology. She joined as a lecturer in aster of Computer Applications in Pad. D. Y. Patil Master of Computer Application 2005. Presently she is working as a Assistant Professor of Master of Computer Applications in G.H. Raisoni college of Engg. and Mgmt. She has published 2 international papers in journals and conferences. Currently he is working on Cloud computing.